Title	Multivariate Statistical Analysis of Japanese VCV Utterances.
Author(s)	Tabata, Koh-ichi; Sakai, Toshiyuki
Citation	音声科学研究 = Studia phonologica (1973), 7: 31-54
Issue Date	1973
URL	http://hdl.handle.net/2433/52596
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

STUDIA PHONOLOGICA VII (1973)

Multivariate Statistical Analysis of Japanese VCV Utterances

Koh-ichi TABATA and Toshiyuki SAKAI

Summary

Considering the amplitude outputs of 20-channel ¼-octave filter analyzer of V_1 CV_2 utterances as the components of 20-dimensional vector, we performed multivariate analysis of variance with four factors—V_1, C, V_2 and speaker; and compared the amounts of the effects of each factor within themselves.

Choosing one of vowels /a, i, u, e, o/ for V_1 and V_2, and one of nasal consonants /m, n, ŋ/ for C, we made all the combinations with them, and five adult males were asked to utter these 75 kinds of words, which were used for the analysis. Then we inspected the relation of the variance ellipsoids of each factor along their principal axes; signified that the notion of direction as well as amount is necessary for explaning the effects of each factor; and compared these analysis with the principal-component analysis. Another thing we investigated by the method of regression estimate was the relation between final vowels and each section of words. Furthermore, we performed similar analysis on the basis of three-dimensional vectors which consist of the formant frequencies extracted from the same materials as above, and compared these with the case of spectra. The results concerning speech sounds are as follows:

(1) Speaker-effect is considerably large, while consonant-effect is not so large. However, the directions of three distributions of these two effects and vowel-effect meet at nearly right angles with each other:

(2) Intensive correlation is seen between vowel and speaker-factor:

(3) In the case of formant frequency, the informations on any factor other than vowel-factor are being decreased as compared with the case of spectrum distribution.

1 Introduction

The reason that we make syllable sequences—consonant-vowel (CV), vowel-consonant-vowel (VCV), etc.—an object of the basic analysis of the speech sounds is that these phonemes—consonant and vowel—are not uttered independently. It is, of course, basically necessary to investigate the characteristics of the phonemes uttered individually. However, it is more actual to investigate the characteristics of each phoneme in the syllable sequences described above, since, par-

Koh-ichi TABATA (田畑孝一) : Assistant Professor, Department of Information Science, Kyoto University.
Toshiyuki SAKAI (坂井利之) : Professor, Department of Information Science, Kyoto University.

ticularly in Japanese, consonants are seldom uttered individually, but uttered respectively in the form of the syllable which accompanies a vowel.

We investigated the correlations of these phonemes—co-articulation—and the individual differences between various speakers, making the syllable sequences of VCV type an object of analysis.

Although formant frequencies have been utilized traditionally for this kind of research, we have been keenly feeling restrictions on its analysis and insufficient points in its way of explaining the co-articulation. The report of Broad and Fertig(1) states that the patterns of influences of various kinds of initial or final consonants on the vowel region of CVC syllables (sequences of consonant-vowel-consonant: Only the vowel /I/ was used) are quantitatively displayed by an analysis of variance (univariate). It has a statistical ground, so that it is quite interesting. The first, second and third formants are still separately dealt with as univariate, respectively, in their report. However, since there is no guarantee that these formants are independent of each other, and since it is not clear how many parts of sound information the formants share, their report is not sufficient to explain co-articulation. Now, in this paper, it has become possible to deal directly with the spectra of consonants as well as those of vowels by introducing a method of multivariate analysis so that we could be released from the restrictions. Furthermore, we investigated, this time, not only the co-articulations but also the influences of speakers. This fact may not be seen in any other research.

We considered the components of spectrum distributions of a VCV word at various time points as the components of multi-dimensional vectors at first; performed multivariate analysis of variance with four factors—speaker, initial vowel, consonant and final vowel; analyzed the characteristics of co-articulations and individualities of speakers; and treated them quantitatively.(2)(3)

As a result of comparing the values of the factor-effects (obtained from the results of this analysis of variance) with the discrimination scores of each factor, notion of "Direction" as well as "Amount" was found to be necessary for explaining these factor-effects for the first time. We also clarified the difference between this analysis and the principal-component analysis.

Others, we made researches for, are the difference between information included in spectra and that included in formants, and the correlation between each section of VCV word and final vowel by the method of the multiple regression theory. We chose nasal sounds for C, in this chapter, in order to make the analyses easy. This is because nasal sound possesses the best stationariness among consonants.

2 ESTABLISHMENT OF EXPERIMENTAL OBJECTS AND FACTORS

Suppose a V_1 C V_2 (vowel-consonant-vowel) word like /ame/. Choosing one of /a,i,u,e,o/ for the initial vowel V_1 and the final vowel V_2, respectively, and one of /m, n, ŋ/ for the consonant C, we make all the combinations with them. Then, 75 kinds of words are equipped. We made, further, 375 words as materials for the analysis by asking 5 adult males to utter these 75 kinds of words in the simplified nonreverberant room. Schematic representation of V_1 C V_2 word by amplitude is seen in Fig. 1. 9 points of stationary or transition parts of every word are chosen and denoted $t=t_1, \ldots, t_9$, in turn, upon visual observation. Presume the spectrum components at each time to be vector components, and vectors corresponding to the time t to be $\boldsymbol{x}(t)$.

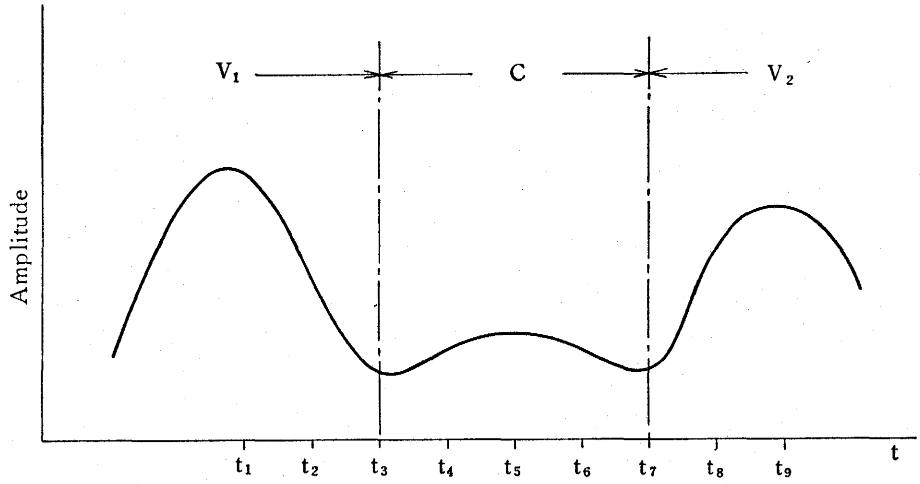

Fig. 1. Schematic representation of V_1CV_2 word and definition of t_i.
$t=t_1$: stationary part of V_1,
$t=t_3$: boundary of V_1-C,
$t=t_5$: stationary part of C,
$t=t_7$: boundary of $C-V_2$,
$t=t_9$: stationary part of V_2.
If i is even, $t_i=(t_{i+1}+t_{i-1})/2$.

We performed multivariate analysis of variance for four-factor (speaker, V_1, C and V_2) design with single observation as Table 1 upon assorting the vectors which correspond to the same time t from all materials.

Table 1. Multivariate analysis of variance for four-factor design with single observation.

Factor	Level	Main effect	No. of levels
A : speaker	A_i	α_i	$i=1 \sim a$ (a=5)
B : V_1	B_j	β_j	$j=1 \sim b$ (b=5)
C : C	C_k	γ_k	$k=1 \sim c$ (c=3)
D : V_2	D_l	δ_l	$l=1 \sim d$ (d=5)

If the outputs of the 20-channel 1/4-octave filter-bank (20 filters whose center frequencies cover 210 up to 5660 Hz) are assumed to be $b_1(t), b_2(t), \dots, b_p(t)$ ($p=20$), in order, $b_1(t), \dots, b_p(t)$ represent the phoneme spectra at time t. After normalizing the square sum of these components at 1, we established p-dimensional vector $\boldsymbol{x}(t)$ by taking the logarithm of its components. Namely, we defined p-dimensional vector

$$\boldsymbol{x}(t) = (x_1(t), \dots, x_p(t)) \quad \text{with} \quad x_i(t) = \log \frac{b_i(t)}{\sqrt{\sum_{i=1}^{p} b_j^2(t)}} \tag{1}$$

which would be used for the analysis. The amplitude outputs of the filter analyzer were AD-converted at every 10 ms, then put into the computer by on-line in real time. The speech spectrum patterns shaded by changing letters were plotted on the line-printer, and then marked either boundary points or stationary parts upon visual observation.

3 LINEAR MODEL AND MULTIVARIATE ANALYSIS OF VARIANCE[4][5][6]

Suppose a linear model, which has the four factors mentioned above, at every $t(=t_1, \dots, t_9)$ for vectors $\boldsymbol{x}(t) = (x_1(t), \dots, x_p(t))$ that represent the spectra (Table 1)

$$\boldsymbol{x}_{ijkl}(t) = \mu(t) + \alpha_i(t) + \beta_j(t) + \gamma_k(t) + \delta_l(t) + \varepsilon_{ijkl}(t), \tag{2}$$

$$(1 \leqq i \leqq a,\ 1 \leqq j \leqq b,\ 1 \leqq k \leqq c,\ 1 \leqq l \leqq d)$$

$$\sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{k=1}^{c} \gamma_k = \sum_{l=1}^{d} \delta_l = 0, \tag{3}$$

$$\varepsilon_{ijkl} \sim N(0, \Lambda) : \text{p-dimensional normal distribution. } \Lambda = \Lambda(t). \tag{4}$$

$\alpha_i(t)$, $\beta_j(t)$, $\gamma_k(t)$ and $\delta_l(t)$ represent the i-th main effect of factor A, j-th of factor B, k-th of factor C and l-th of factor D, respectively. We determine $\mu(t)$ in order to satisfy the condition given by Eq. (3), and assume $\varepsilon_{ijkl}(t)$ to be independently distributed according to the p-dimensional normal distribution.

Letting " ' " represent transposed matrix, the breakdown of the total variance $Q(p \times p)$ (matrix of sums of squares and cross products) becomes as in Eq. (5).

(We disregard t as long as there seems to be no misunderstanding.)

$$Q = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{d} (\boldsymbol{x}_{ijkl} - \boldsymbol{x}_{\cdots\cdot})'(\boldsymbol{x}_{ijkl} - \boldsymbol{x}_{\cdots\cdot}) = Q_1 + Q_2 + Q_3 + Q_4 + R, \tag{5}$$

where

$$Q_1 = \sum_{i=1}^{a} (\boldsymbol{x}_{i\cdots} - \boldsymbol{x}_{\cdots\cdot})'(\boldsymbol{x}_{i\cdots} - \boldsymbol{x}_{\cdots\cdot})$$

$$Q_2 = \sum_{j=1}^{b} (\boldsymbol{x}_{\cdot j\cdot\cdot} - \boldsymbol{x}_{\cdots\cdot})'(\boldsymbol{x}_{\cdot j\cdot\cdot} - \boldsymbol{x}_{\cdots\cdot})$$

$$Q_3 = \sum_{k=1}^{c} (\boldsymbol{x}_{\cdot\cdot k\cdot} - \boldsymbol{x}_{\cdots\cdot})'(\boldsymbol{x}_{\cdot\cdot k\cdot} - \boldsymbol{x}_{\cdots\cdot})$$

$$Q_4 = \sum_{l=1}^{d} (\boldsymbol{x}_{\cdots l} - \boldsymbol{x}_{\cdots\cdot})'(\boldsymbol{x}_{\cdots l} - \boldsymbol{x}_{\cdots\cdot})$$

$$R=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i\ldots}-\boldsymbol{x}_{.j..}-\boldsymbol{x}_{..k.}-\boldsymbol{x}_{\ldots l}+3\boldsymbol{x}_{\ldots.})'$$
$$\cdot(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i\ldots}-\boldsymbol{x}_{.j..}-\boldsymbol{x}_{..k.}-\boldsymbol{x}_{\ldots l}+3\boldsymbol{x}_{\ldots.}),$$

and

$$\boldsymbol{x}_{\ldots.}=\frac{1}{abcd}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\boldsymbol{x}_{ijkl},\quad \boldsymbol{x}_{i\ldots}=\frac{1}{bcd}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\boldsymbol{x}_{ijkl},$$
$$\boldsymbol{x}_{.j..}=\frac{1}{acd}\sum_{i=1}^{a}\sum_{k=1}^{c}\sum_{l=1}^{d}\boldsymbol{x}_{ijkl},\quad \boldsymbol{x}_{..k.}=\frac{1}{abd}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{l=1}^{d}\boldsymbol{x}_{ijkl},$$
$$\boldsymbol{x}_{\ldots l}=\frac{1}{abc}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\boldsymbol{x}_{ijkl}.$$

Q_1, Q_2, Q_3, and Q_4, represent the variances (Matrices of sums of squares and cross products) corresponding to factors A, B, C and D, respectively, and R (Matrix of sums of squares and cross products) does the residual.

Now, we shall develop a test of the hypothesis for factor A at time t ($=t_1, \ldots, t_9$) that all the effects of A_i are equal (there is no effect of A);

$$H_A(t): \alpha_1(t)=\ldots..=\alpha_a(t)=0. \tag{6}$$

We can test the hypothesis since it is possible to prove that the likelihood ratio criterion

$$\nu=\left\{n-l_2-\frac{1}{2}(p+l_1+1)\right\}\log\frac{|Q_1+R|}{|R|} \tag{7}$$

is distributed asymptotically according to χ^2-distribution with pl_1 degrees of freedom under the condition—$n=a\cdot b\cdot c\cdot d$, $l_1=a-1$, $l_1+l_2=a+b+c+d-3$, $n-a-b-c-d+3\geqq p$—when n is sufficiently large. (See Appendix A.) We obtained test criterion ν for each factor at $t=t_1, \ldots, t_9$ based on the materials mentioned above (Fig. 1). As the degree of freedom of the factors are different from each other (See Table 2)

Table 2. Degrees of freedom of Q_i ;
degrees of freedom of R is $n-a-b-c-d+3$ ($n=abcd$)

Factor	A	B	C	D
Q_i	Q_1	Q_2	Q_3	Q_4
l_1	$a-1$	$b-1$	$c-1$	$d-1$
pl_1	$p(a-1)$	$p(b-1)$	$p(c-1)$	$p(d-1)$
l_1+l_2	$a+b+c+d-3$	$a+b+c+d-3$	$a+b+c+d-3$	$a+b+c+d-3$

and so are the value of $\chi^2-1\%$ significant level different from each other, we signified the normalized criterion

$$\nu'=\frac{\nu}{(\text{Value of } \chi^2-1\% \text{ significant level corresponding to the degrees of freedom of } \nu)} \tag{8}$$

in Fig. 2. From Fig. 2, it can be said that

(1) The effect (main effect) of the speaker-factor is the largest among four factors at the stationary part of nasal consonant:

(2) The effect of the vowel-factor is the largest at the stationary part of vowel among four factors.

(3) Although a result does not insist that an effect of V_2 (V_1) is observed at the stationary part of V_1 (V_2) ($\nu'<1$), it may be necessary to scheme a more accurate experiment in order to assert it.

(4) At the stationary part of V_1 (V_2), the effect of factor C is observed but it is smaller than that of speaker-factor.

(5) The effect of factor C is maximum at the part of C among all sections. The effect of C at V_2 is larger than that of C at V_1.

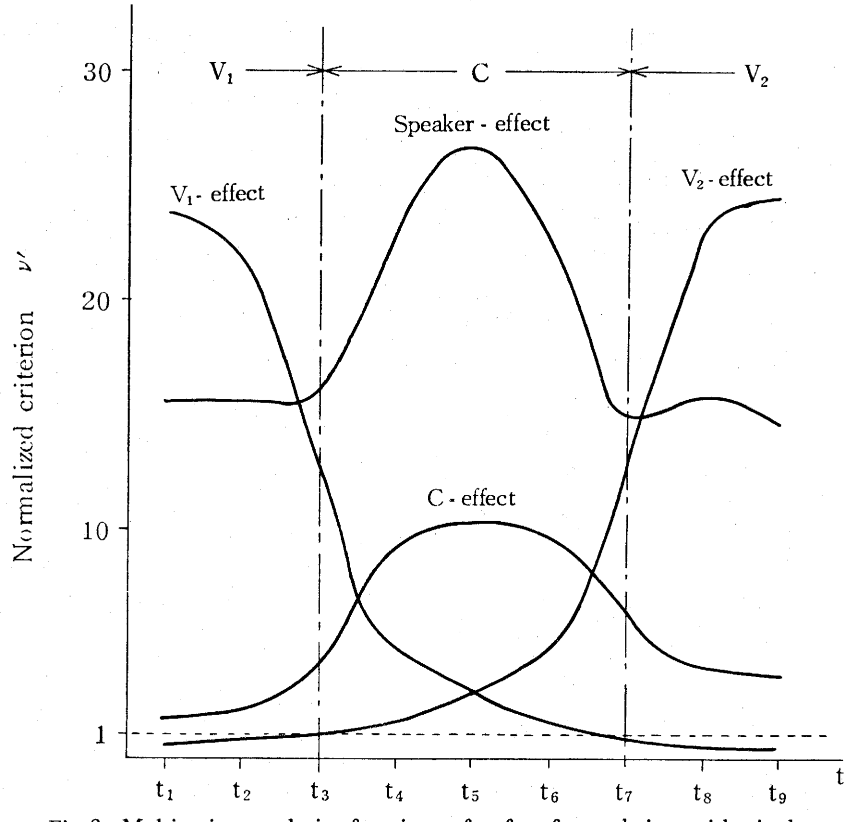

Fig. 2. Multivariate analysis of variance for four-factor design with single observation.

4 RELATION BETWEEN NORMALIZED CRITERION ν' AND DISCRIMINATION SCORE

The larger normalized criterion ν' for factor A is than 1, the more α_i ($i=1\sim a$) differ greatly from each other. Hence, when a certain $\boldsymbol{x}$ (one of $\boldsymbol{x}_{ijkl}(t)$) is given, it may be easier to discriminate in which category (that corresponds to the level A_i in the case of factor A) $\boldsymbol{x}$ belongs. We tried discrimination in each case of all factors at $t(=t_1, \ldots, t_9)$ by using the distance which is based on quadratic form in order to investigate the relation between ν' and discrimination. For example, we calculated the mean vector $\boldsymbol{x}_{i\cdots}(t)$ and the sample covariance matrix $S_i(t)$ for the category which corresponds to each level A $(i=1,\ldots,a)$ of factor A at time t, using $\boldsymbol{x}$'s which belong in that category.

$$\boldsymbol{x}_{i\cdots}(t)=\frac{1}{bcd}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\boldsymbol{x}_{ijkl}(t) \tag{9}$$

$$S_i(t)=\frac{1}{bcd}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(\boldsymbol{x}_{ijkl}(t)-\boldsymbol{x}_{i\cdots}(t))'(\boldsymbol{x}_{ijkl}(t)-\boldsymbol{x}_{i\cdots}(t)) \tag{10}$$

Assume, now, that we discriminate that given $\boldsymbol{x}(t)$ (one of $\boldsymbol{x}_{ijkl}(t)$) belongs in the

Table 3. Discrimination scores in each factor (%).

Factor	No. of categories	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
A : speaker	5	100	99.8	99.8	100	100	99.8	99.8	99.5	98.7
B : V_1	5	100	100	96.5	79.5	69.9	64.5	55.7	52.8	53.1
C : C	3	69.6	73.1	85.1	95.2	95.5	95.5	89.1	80.3	77.9
D : V_2	5	54.7	53.3	58.1	64.0	72.0	82.7	96.0	100	100

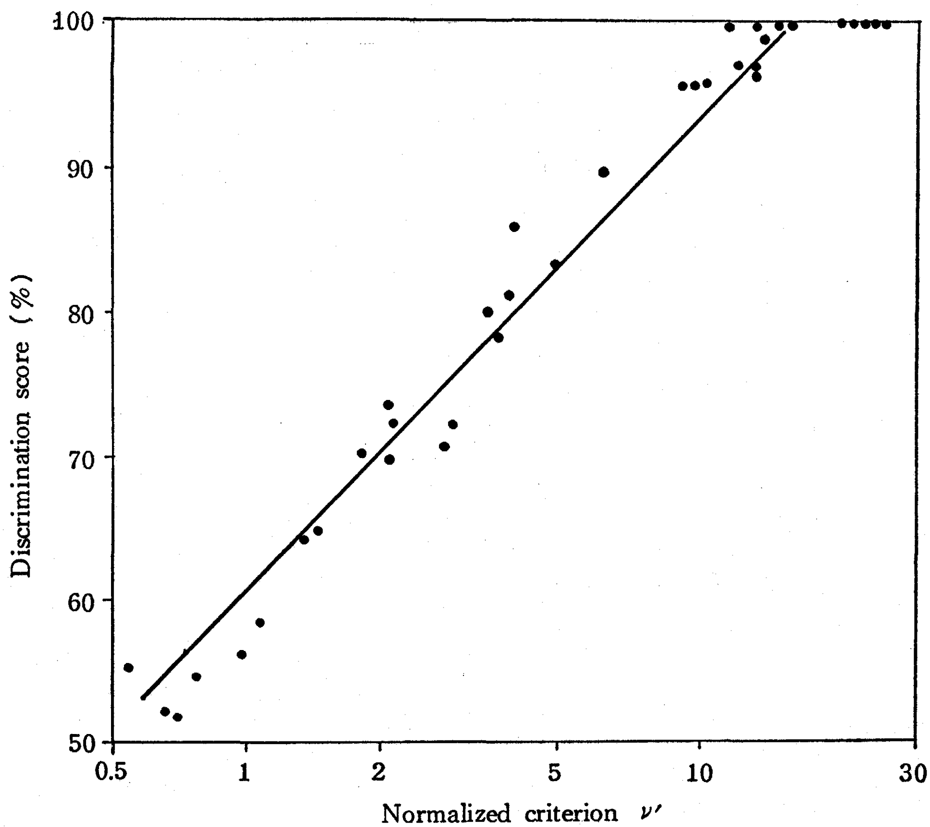

Fig. 3. Discrimination scores vs. normalized criterion ν'.

category A_i where "i" makes

$$(\boldsymbol{x}(t)-\boldsymbol{x}_{i\cdots}(t))S_i^{-1}(t)(\boldsymbol{x}(t)-\boldsymbol{x}_{i\cdots}(t))'+\log|S_i(t)| \tag{11}$$

minimum among $i=1,\ldots,a$.

If every category has the normal distribution and frequencies of occurrence of each category of $A_1 \sim A_a$ are the same, the notion of discrimination of Eq. (11) corresponds that "The category, in which the probability of the occurrence of given $\boldsymbol{x}(t)$ is the largest, is A_i." (Bayes' theorem.) The discrimination scores for each time and each factor obtained by this judgement are presented in Table 3. The discrimination scores of the factors, whose effects were the largest in analysis of variance at each time, are nearly 100%, and that of the factors, whose effects are the second largest, are not so unexpected, too. Although C-effects at the stationary part of nasal ($t=t_5$) and at the boundary between nasal and vowel ($t=t_7$) are considerably small comparing with the largest effect, we should notice that the discrimination scores of the nasals are 95.5% and 89.1%, respectively.

We will speak further of the fact that there is the relation described in Fig. 3 between the discrimination scores obtained here and the normalized criterion ν'. Thus, it is found that ν' has close relations with the discrimination score though ν' means originally the statistics for testing the hypothesis.

5 GEOMETRIC REPRESENTATION OF MULTI-FACTOR DISTRIBUTIONS

As signified in Section 4, it is supposed that the directions of the distributions of variances of each factor may be different from each other by the reason that the discrimination scores of the second and the third largest factors (which possess considerably small value of effect as compared with the largest effect) do not become worse. As it deals with ratio $|Q_i+R|/|R|$, that is, the ratio of the variance of each factor to the residual variance, in analysis of variance as described by Eq. (7), we can observe only, so to speak, the relative largeness of the distributions of each factor.

Therefore, we scheme geometric interpretation of the distribution as follows in order to clarify the relation between the directions of distributions of each factor.

At first, if we generally let μ, Λ be the expected vector and the covariance matrix of the probability vector of $\boldsymbol{x}$ $(1\times p)$, respectively, we may think that

$$(\boldsymbol{x}-\mu)\Lambda^{-1}(\boldsymbol{x}-\mu)'=p+2 \tag{12}$$

expresses geometrically the pattern of the variance of $\boldsymbol{x}$, where Eq. (12) represent a concentration ellipsoid[4] for $\boldsymbol{x}$.

In this paper, we are going to signify the variance of $\boldsymbol{x}$ by utilizing the following ellipsoid (13) which is similar to the above ellipsoid (12) (similarity ratio $1/\sqrt{p+2}$).

$$(\boldsymbol{x}-\hat{\mu})\hat{\Lambda}^{-1}(\boldsymbol{x}-\hat{\mu})'=1. \tag{13}$$

Where $\hat{\mu}$ and $\hat{\Lambda}$ are the maximum likelihood estimates of μ and Λ, respectively.

(Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be sample vectors, then

$$\hat{\mu} = \boldsymbol{x}. = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i, \text{ and } \hat{\Lambda} = \frac{1}{n}Q = \frac{1}{n}\sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{x}.)'(\boldsymbol{x}_i - \boldsymbol{x}.).)$$

The reason is that (13) is easy to understand numerically because the distance between center $\hat{\mu}$ and the point of intersection (produced by the ellipsoid represented by Equation (13) and the principal axis—eigenvector—corresponding to eigenvalue of $\hat{\Lambda}$) is just $\sqrt{\sigma_i}$ if one of eigenvalues of $\hat{\Lambda}$ is σ_i (i=1~p). (See Appendix B).

Next, as shown in Appendix C, Q_1+R can be thought to express the variance of factor A, and $\frac{1}{n}(Q_1+R)$ can be also thought to express the covariance matrix of factor A.

We, further, project the variance Q_1+R onto the new vector space obtained by normalizing the original vector space by the residual R.

Suppose the nonsingular linear transformation

$$\boldsymbol{x} \rightarrow \tilde{\boldsymbol{x}} = \boldsymbol{x}(R/n)^{-\frac{1}{2}}, \tag{14}$$

where R is the residual variance, and n (=abcd) is sample size. ($\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$ are vectors in the original space and the new space, respectively.)

Then, the covariance matrix $\frac{1}{n}(Q_1+R)$ is transformed to

$$\frac{1}{n}(\tilde{Q}_1+\tilde{R}) = R^{-\frac{1}{2}}(Q_1+R)R^{-\frac{1}{2}}. \tag{15}$$

(See Appendix D.)

So that R itself becomes

$$\frac{1}{n}\tilde{R} = R^{-\frac{1}{2}}RR^{-\frac{1}{2}} = I_p \tag{16}$$

(Note that $(Q_1+R)R^{-1}$ is also considered to be another normalization, but it is not always symmetric matrix, so that it is unsuitable for geometric expression.)

We will continue to discuss Eq. (15). As the residual matrix R is symmetric and positive difinite (provided, $n-a-b-c-d+3 \geqq p$. See Eq. (A.7) of Appendix A), $R^{\frac{1}{2}}$ exists. Where $R^{\frac{1}{2}}R^{\frac{1}{2}}=R$ and $R^{-\frac{1}{2}}$ is the inverse matrix of $R^{\frac{1}{2}}$. Matrix Q_1 is symmetric and its rank is $a-1$ (provided $a-1 \leqq p$), and $R^{\frac{1}{2}}$ is also symmetric and real nonsingular. Accordingly $R^{-\frac{1}{2}}Q_1R^{-\frac{1}{2}}$ becomes symmetric and its rank is $a-1$. Hence, it follows that the eigenvalues of

$$R^{-\frac{1}{2}}Q_1R^{-\frac{1}{2}}\boldsymbol{z}' = \sigma\boldsymbol{z}'$$

are $\sigma_1 > \sigma_2 \ldots > \sigma_{a-1} > 0$ (with probability 1) and $\sigma_a = \ldots = \sigma_p = 0$.

Then the eigen-values of

$$R^{-\frac{1}{2}}(Q_1+R)R^{-\frac{1}{2}}\boldsymbol{a}' = \lambda\boldsymbol{a}' \tag{17}$$

become $\lambda_1 > \lambda_2 > \ldots > \lambda_{a-1} > 1$, $\lambda_a = \ldots = \lambda_p = 1$. Because, $\lambda = \sigma + 1$ is shown from the fact

$$\lambda\boldsymbol{a}' = [R^{-\frac{1}{2}}Q_1R^{-\frac{1}{2}} + I_p]\boldsymbol{a}' = R^{-\frac{1}{2}}Q_1R^{-\frac{1}{2}}\boldsymbol{a}' + I_p\boldsymbol{a}' = \sigma\boldsymbol{a}' + \boldsymbol{a}' = (\sigma+1)\boldsymbol{a}'.$$

Also,

$$\frac{|Q_1+R|}{|R|}=|R^{-\frac{1}{2}}(Q_1+R)R^{-\frac{1}{2}}|=\lambda_1\cdot\lambda_2\ldots\lambda_{s-1}. \tag{18}$$

When λ_1 is named the maximum eigenvalue, and the eigenvector $\boldsymbol{a}_1$ corresponding to it is provisionally named the first principal axis of the ellipsoid which is expressed by

$$\tilde{\boldsymbol{x}}\left[\frac{1}{n}(\tilde{Q}_1+\tilde{R})\right]^{-1}\tilde{\boldsymbol{x}}'=\tilde{\boldsymbol{x}}\left[R^{-\frac{1}{2}}(Q_1+R)R^{-\frac{1}{2}}\right]^{-1}\tilde{\boldsymbol{x}}'=1 \tag{19}$$

and which represents the variance of factor A, $\sqrt{\lambda_1}$ and $\boldsymbol{a}_1$ are considered to be the amount and the direction of the substantial proportion of the variance of factor A, respectively.

At $t=t_7$ (at the C-V_2 boundary), the maximum eigenvalues and the first principal axes of speaker-factor $(\tilde{Q}_1+\tilde{R})$, C-factor $(\tilde{Q}_3+\tilde{R})$ and V_2-factor $(\tilde{Q}_4+\tilde{R})$ were computed, then the results became $(2.4^2, \boldsymbol{a}_1)$, $(1.4^2, \boldsymbol{c}_1)$ and $(2.5^2, \boldsymbol{d}_1)$, respectively. These are illustrated in Fig. 4.

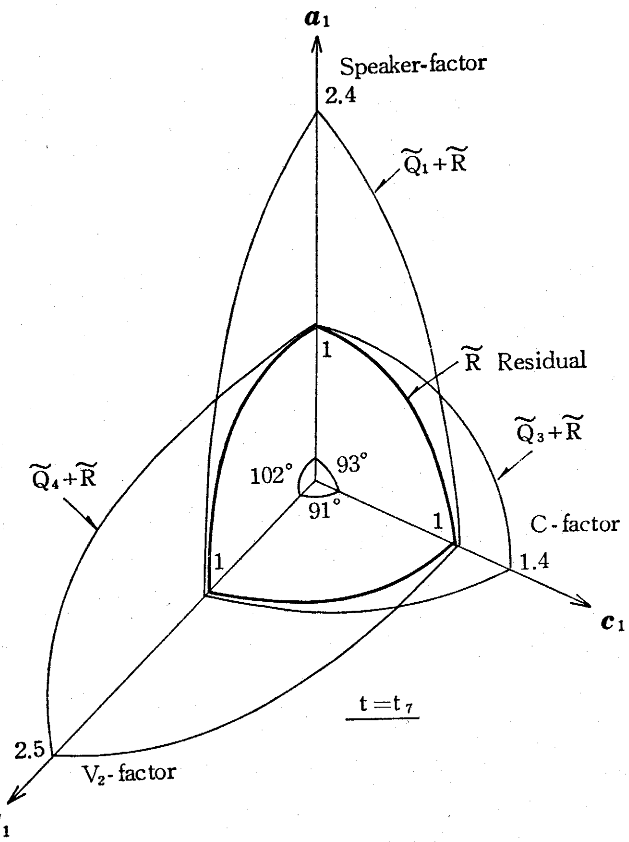

Fig. 4. Geometric representation of multi-factor distributions $(t=t_7)$.

The intersections, made by three ellipsoids (that are represented by $\tilde{\boldsymbol{x}}\left[\frac{1}{n}(\tilde{Q}_i+\tilde{R})\right]^{-1}\tilde{\boldsymbol{x}}'=1$ $(i=1, 3, 4)$) and three planes that are determined by $\boldsymbol{a}_1$ & $\boldsymbol{c}_1$, $\boldsymbol{c}_1$ &

$\boldsymbol{d}_1$ and $\boldsymbol{d}_1$ & $\boldsymbol{a}_1$, are also illustrated in the figure. (How to obtain the intersections is in Appendix B.) The angle θ between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined by $\theta = \cos^{-1}\frac{\boldsymbol{xy}'}{(\boldsymbol{xx}')^{\frac{1}{2}}(\boldsymbol{yy}')^{\frac{1}{2}}}$. The angle between $\boldsymbol{a}_1$ and $\boldsymbol{c}_1$ is 93°, and so on. Besides,

$$\tilde{\boldsymbol{x}}\left(\frac{1}{n}\tilde{\mathrm{R}}\right)^{-1}\tilde{\boldsymbol{x}}' = \tilde{\boldsymbol{x}}\mathrm{I}_p\tilde{\boldsymbol{x}}' = \tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}' = 1$$

for the resiual R. From this description, it can be realized that the principal axes of the variances of these three factors meet nearly at right angles with each other: Similarly, the distributional relations between the speaker-factor ($\tilde{\mathrm{Q}}_1 + \tilde{\mathrm{R}}$) and C-factor ($\tilde{\mathrm{Q}}_3 + \tilde{\mathrm{R}}$) at $t = t_5$ (that is the stationary point of C), and that between speaker-factor ($\tilde{\mathrm{Q}}_1 + \tilde{\mathrm{R}}$) and V_2-factor ($\tilde{\mathrm{Q}}_4 + \tilde{\mathrm{R}}$) at $t = t_9$ (that is the stationary point of V_2) are as Fig. 5. From this figure, we can understand that the discrimination score does not decrease since the directions of the variances are different from each other even if the amount of variance is slight.

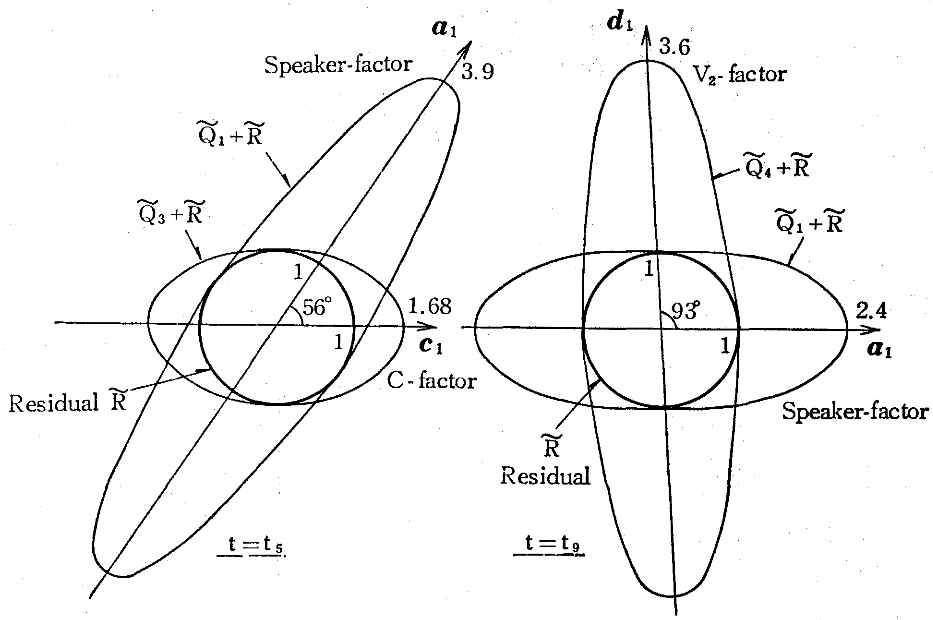

Fig. 5. Geometric representation of multi-factor distributions ($t = t_5$, $t = t_9$).

6 Comparison with Principal-Component Analysis

Klein, Plomp and Pols[7] express vowels by using the first four principal components of a principal-component analysis and argue the distribution of vowels and speaker's individualities, regarding the amplitude outputs of 18-channel 1/3-octave filters as the components of 18-dimensional vector. (They dealt with 600 utterances—12 kinds of vowels by pronounced by 50 male speakers.) However, their explanations are not direct because they consider projection of vowels on the plane determined by the principal axes of a principal-component analysis which make neither the vowel-factor nor the speaker-factor maximum.

We observed the following in order to clarify how each factor is expressed by a principal-component analysis.

Total variance Q of Eq. (5) is exactly the same as the sample covariance matrix S which is used in a principal-component analysis.[8][9] Namely, Q= nS, where n is the number of $\boldsymbol{x}$'s. As the eigenvectors of S are arranged in order of largeness of the eigenvalues corresponding to them, and are named $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots$, respectively, the inner product by $\dot{\boldsymbol{x}}$ $(=\boldsymbol{x}_{\mathrm{ijkl}}(\mathrm{t})-\boldsymbol{x}_{\ldots.}(\mathrm{t}))$ and $\boldsymbol{e}_{\mathrm{m}}$ makes the m-th principal component of $\dot{\boldsymbol{x}}$ $(\boldsymbol{e}_i\boldsymbol{e}_i'=1,\ \boldsymbol{e}_i\boldsymbol{e}_j'=0,\ i\neq j)$. If the 2-dimensional vector made by the first and the second principal components is represented by $\bar{\boldsymbol{x}}$, $\bar{\boldsymbol{x}}$ denotes the orthogonal projection of $\dot{\boldsymbol{x}}$ on $\boldsymbol{e}_1-\boldsymbol{e}_2$ plane. (The variance explained by the first and second principal compenent in this case is 83% of the total variance.) The breakdown of the total variance $\bar{\mathrm{Q}}$ by using $\bar{\boldsymbol{x}}$ similarly to Eq. (5) is as

$$\bar{\mathrm{Q}}=\bar{\mathrm{Q}}_1+\bar{\mathrm{Q}}_2+\bar{\mathrm{Q}}_3+\bar{\mathrm{Q}}_4+\bar{\mathrm{R}}, \tag{20}$$

where $\bar{\mathrm{Q}}_1$, $\bar{\mathrm{R}}$ have the same meaning as Q_1 and R in Eq. (5) do, and p=2. In this case, $\bar{\boldsymbol{x}}_{\ldots.}=\frac{1}{\mathrm{abcd}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\bar{\boldsymbol{x}}_{\mathrm{ijkl}}=0$, because

$$\frac{1}{\mathrm{abcd}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}\dot{\boldsymbol{x}}_{\mathrm{ijkl}}=\frac{1}{\mathrm{abcd}}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(\boldsymbol{x}_{\mathrm{ijkl}}-\boldsymbol{x}_{\ldots.})=0.$$

Similar to the preceding section, $\bar{\mathrm{R}}$ represents the residual variance. Variance of factor A, for example, can be regarded as $\bar{\mathrm{Q}}_1+\bar{\mathrm{R}}$. And let $\bar{\boldsymbol{x}}\left(\frac{\bar{\mathrm{R}}}{\mathrm{n}}\right)^{-1}\bar{\boldsymbol{x}}'=1$, $\bar{\boldsymbol{x}}\left[\frac{1}{\mathrm{n}}(\bar{\mathrm{Q}}_1+\bar{\mathrm{R}})\right]^{-1}\bar{\boldsymbol{x}}'=1$ be the ellipsoids which denote these variances, respectively.

The illustrations of the ellipsoids are drawn in Fig. 6.

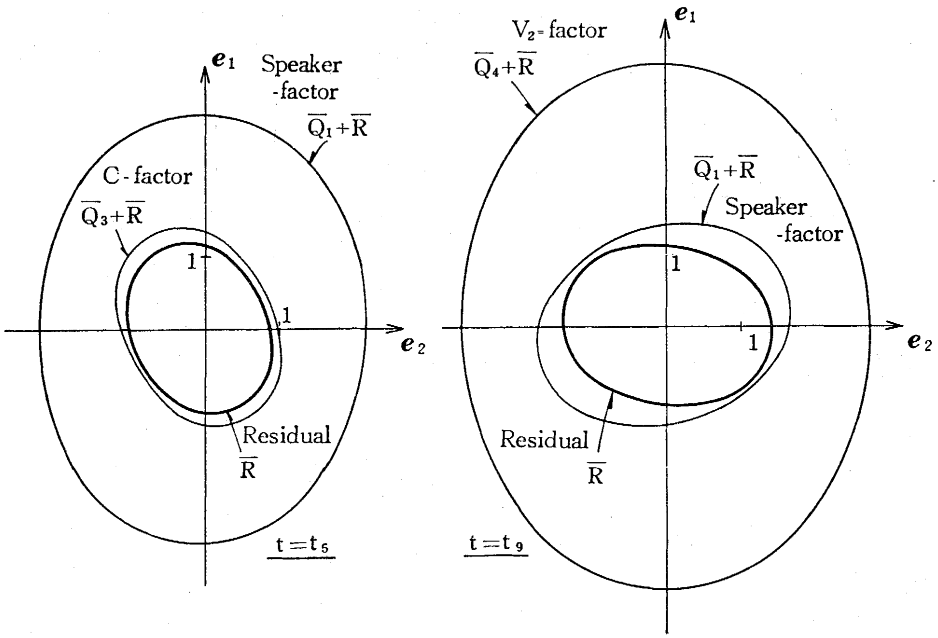

Fig. 6. Principal-component analysis $(\mathrm{t}=\mathrm{t}_5,\ \mathrm{t}=\mathrm{t}_9)$.

One of that denotes speaker-factor ($\bar{Q}_1+\bar{R}$), factor C($\bar{Q}_3+\bar{R}$), and residual $\bar{R}$ at $t=t_5$, and the other denotes speaker-factor ($\bar{Q}_3+\bar{R}$), factor $V_2(\bar{Q}_4+\bar{R})$ and residual $\bar{R}$ at $t=t_9$. Comparing Fig. 6 with Fig. 5, it is understood that the factors which have the second and less largest effects are lost in the variance of the factor which has the largest effect, and that the ratios of these variances to the residual variance are not so favorable.

To be brief as to characteristics of the principal-component analysis in this study, what is mainly explained is about the superior factor, but it is difficult for the inferior factors to be explained sufficiently. It is just like, "The weak become the victim of the strong."

7 Discussion about the Residual

By the model of Eq. (2), it is presumed that Eq. (4)

$$\varepsilon_{ijkl}=\boldsymbol{x}_{ijkl}-\mu-\alpha_i-\beta_j-\gamma_k-\delta_l \tag{21}$$

has the p-dimensional normal distribution $N(0, \Lambda)$. If μ, α_i, β_j, γ_k and δ_l are replaced by maximum likelihood estimates $\boldsymbol{x}_{\dots\dots}$, $\boldsymbol{x}_{i\dots}-\boldsymbol{x}_{\dots\dots}$, ..., and $\boldsymbol{x}_{\dots l}-\boldsymbol{x}_{\dots\dots}$ (Eq. (A.2) of Appendix A), then χ_{ijkl} replacing ε_{ijkl},

$$\chi_{ijkl}=\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i\dots}-\boldsymbol{x}_{.j..}-\boldsymbol{x}_{..k.}-\boldsymbol{x}_{\dots l}+3\boldsymbol{x}_{\dots\dots}, \tag{22}$$

is desired to be distributed according to $N(0, \Lambda)$. Now, we are going to investigate the distributional pattern of each coordinate component upon transforming coordinate in order that $\hat{\Lambda}=(1/n)R$ becomes a diagonal matrix. We utilized $\hat{\Lambda}$, maximum likelihood estimate of Λ, instead of Λ which is unknown.

Arrange the eigenvalues of $\hat{\Lambda}$ in order of largeness; denote them by $\sigma_1^2,\dots,\sigma_p^2$; and let $\tau_1,\dots,\tau_p$ represent the eigenvectors which correspond to them.

If the m-th component of

$$\boldsymbol{y}_{ijkl}=(\chi_{ijkl}\,\tau_1',\dots,\chi_{ijkl}\,\tau_p')$$

is supposed to be $y_{m,ijkl}$. These components are uncorrelated with each other. If $\chi_{ijkl}\sim N(0, \hat{\Lambda})$,

$y_{m,ijkl}/\sigma_m\sim N(0, 1)$: Single dimensional standard normal distribution.

is obtained, since

$$y_{m,ijkl}=\chi_{ijkl}\,\tau_m'\sim N(0, \tau_m\,\hat{\Lambda}\tau_m')=N(0, \sigma^2_m).$$

The observed cumulative frequency distribution $S_N(X)$ of $y_{m,ijkl}/\sigma_m$, plotted on normal probability coordinates, becomes, for example, as Fig. 7, when $t=t_7$ and $m=1$. The distribution shown in Fig. 7 is well-approximated by the normal distribution.

To examine whether these observed cumulative frequency were obtained from the population having the normal distribution or not, we will utilized KS-test (Kolmogorov-Smirnov one-sample test). (See Appendix E.)

Let $F_0(X)$ denote the cumulative frequency distribution function of the standard normal distribution $N(0, 1)$, that is, the straight line of Fig. 7, and the band of $F_0(X)\pm D_\alpha$ (from the mathematical table of KS-test, the critical value $D_\alpha=$

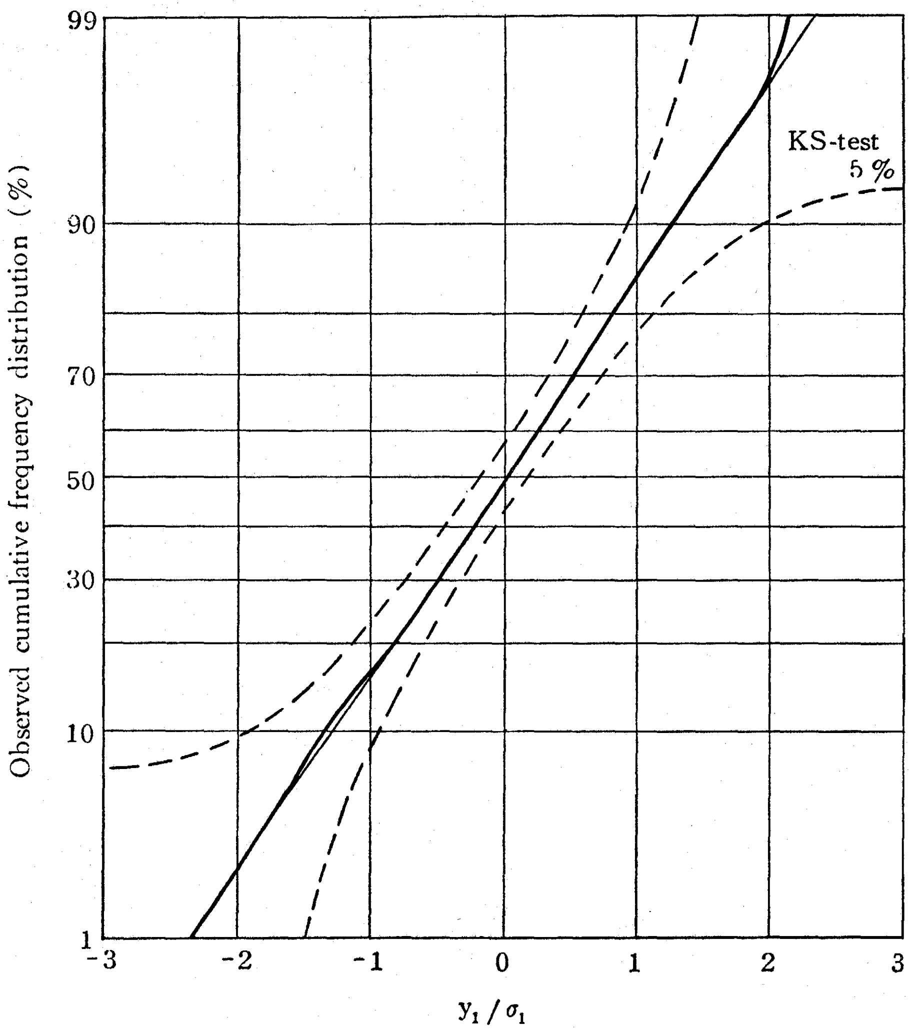

Fig. 7. Observed cumulative frequency distribution of the first component of the residual. ($t=t_7$).

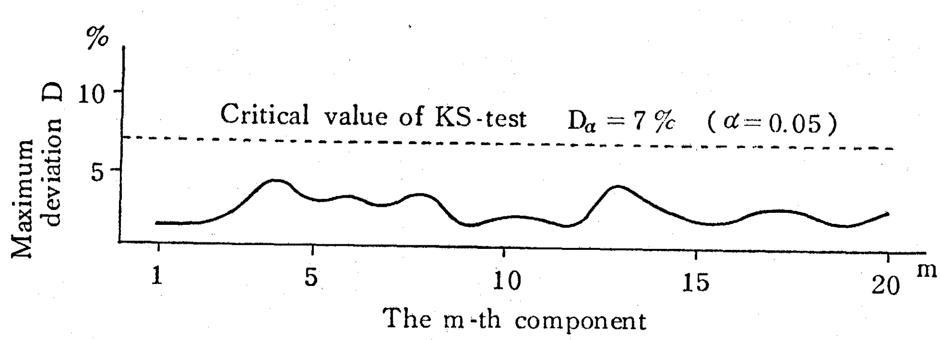

Fig. 8. Maximum deviation D of each component and the critical value of KS-test.

7% when N=375, level of significance α=5%) is between two dotted lines in the same figure. The fact that the line denoting $S_N(X)$ does not cross beyond the dotted lines reveals that $y_{m,1jkl}/\sigma_m$ has the standard normal distribution N(0, 1). Fig. 8 shows the result of obtaining D for each $y_{m,1jkl}/\sigma_m$ (m=1~p) at $t=t_7$ and of comparing this with significant-level value D_α. Each value of D is smaller than D_α so that all of these seems to have univariate normal distributions. Although $\boldsymbol{y}_{1jkl}$ is not always distributed according to p-dimensional normal distribution even if every component of $\boldsymbol{y}_{1jkl}$ has the univariate normal distribution, the assumption on residual $\boldsymbol{\chi}_{1jkl}$ may be almost appropriate, considering from these results.

8 Analysis by Formant Frequency

We analyzed the same speech materials in a way similar to Section 3, considering the first, the second and the third formant frequencies (represented by F_1, F_2 and F_3, respectively) as three-dimensional vectors (p=3). "Formant Frequency Extraction by a inverse Filter and Moment Calculation"[10] reported by Nakatsui and Suzuki was used for extracting the formants from the spectra, obtained through the 1/4-octave filters mentioned before. To describe the accuracy of this measurement, the error of extracting the frequencies in the case of the synthetic speech sound is less than 3.6%.[11] However, we sometimes relied on visual inspectation since it is not always easy to extract in the actual case which contains nasalized vowels.

Formant frequencies obtained in the same data mentioned above at t ($=t_1$, t_3, t_7, t_9), and we carried out multivariate analysis at these times. (Only /a,u,o/

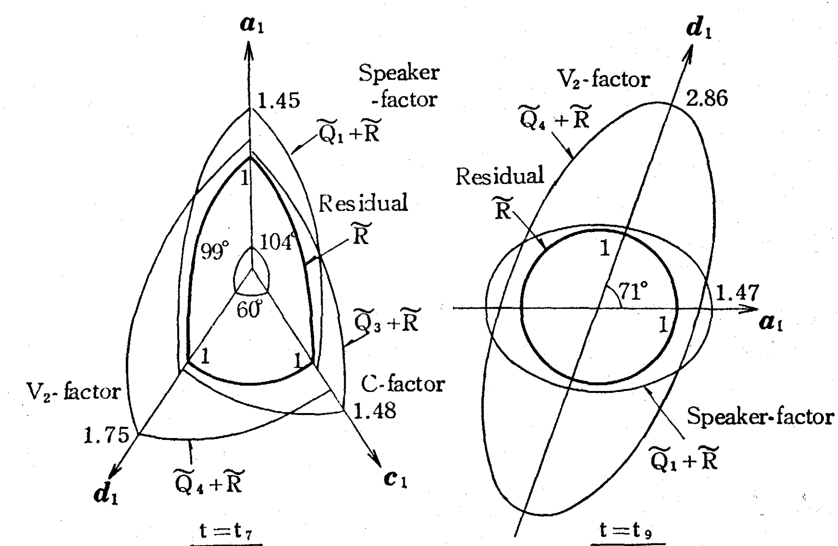

Fig. 9. Geometric representation of multi-factor distributions in the case of the formant frequencies ($t=t_7$, $t=t_9$).

are used as objects for V_2 in the case of formant. V_1 & C are the same as before.) The illustration of the ellipsoids, that represent variance of factors in the meaning of Section 5 when $t=t_7$, t_9, is as Fig. 9.

Meanwhile, Table 4 gives the discrimination scores obtained by formant in the meaning of Section 4. By comparing Fig. 9 with Fig. 4 and Fig. 5, it is possible to explain the reason that the discrimination score for each factor in the case of using formants becomes worse than in the case of using spectra.

Generally, it can be said that formant is rather a "vocalic factor." The informations on any factor other than vowel-factor are being decreased as compared with the case of spectrum distribution.

Table 4. Discrimination scores in each factor (%) ; by formant frequencies.

Factor	No. of categories	t_1	t_3	t_7	t_9
A : speaker	5	43.1	45.3	56.0	46.2
B : V_1	5	96.0	84.0	30.7	24.0
C : C	3	37.3	49.3	64.4	49.3
D : V_2	3	37.3	39.6	77.3	99.6

9 REGRESSION ESTIMATE BY USING FINAL VOWELS AND ANALYSIS OF VARIANCE

In the analysis up to the section above, the spectrum sections at different time points ($t=t_1 \sim t_9$) have been treated as independent sections of each other. Here, one aspect of the correlation between the spectrum sections at various time points in VCV words is examined.

Suppose the equations of multivariate linear regression at each time t in order to analyze similarly to Section 3 after eliminating the influence of the final vowel from each spectrum section.

Let $\boldsymbol{x}_\alpha(t_9)$ (a stationary part of the final vowel) represent a known vector, and $\boldsymbol{x}_\alpha(t)$ observation vector;

presume

$$\boldsymbol{x}_\alpha(t)\,(1\times p)=\boldsymbol{v}(t)+\boldsymbol{x}_\alpha(t_9)\cdot B(t)+\varepsilon_\alpha(t)\ (t=t_1,\ldots,t_8) \tag{23}$$

$$\varepsilon_\alpha(t)\sim N(0,\Lambda),\ B(t):\ p\times p\ \text{matrix};\ \Lambda=\Lambda(t);$$

and estimate $\boldsymbol{v}(t)$ and B(t), where $\boldsymbol{x}_\alpha$ is one of $\boldsymbol{x}_{ijkl}$ and $\alpha=1,\ldots,n$. Only /a,u,o/ are used as V_2 in this section. (V_1, C are the same as in Section 3). Hence, $n=a\cdot b\cdot c\cdot d=5\times5\times3\times3=225$. Supposing $X'(p\times n)=(\boldsymbol{x}_1'(t),\ldots,\boldsymbol{x}_n'(t))$, $Z'(\overline{p+1}\times n)=(\boldsymbol{z}_1',\boldsymbol{z}_2',\ldots,\boldsymbol{z}_n')$, and $\boldsymbol{z}_\alpha(1\times\overline{p+1})=(1,\boldsymbol{x}_\alpha(t_9))$, $\hat{\boldsymbol{v}}(t)$ and $\hat{B}(t)$, maximum likelihood estimates of $\boldsymbol{v}(t)$ and B(t), respectively, are obtained from Eq. (23).

$$\begin{bmatrix}\hat{\boldsymbol{v}}(t)\\ \hat{B}(t)\end{bmatrix}=(Z'Z)^{-1}\ (Z'X). \tag{24}$$

By using these $\hat{v}(t)$ and $\hat{B}(t)$, estimate $x_\alpha(t)$ from $x_\alpha(t_9)$ in each word, one of the same VCV words used in Eq. (23), and let $\hat{x}_\alpha(t)$ be the estimate of $x_\alpha(t)$. The results of the multivariate analysis of variance for the difference $\overset{\blacktriangle}{x}_\alpha(t)$;

$$\overset{\blacktriangle}{x}_\alpha(t) = x_\alpha(t) - \hat{x}_\alpha(t) = x_\alpha(t) - \hat{v}(t) - x_\alpha(t_9) \cdot \hat{B}(t)$$

(similarly to Section 3) is drawn as Fig. 10. (Dotted lines of the same figure indicate the analysis for $x_\alpha(t)$ itself.)

As a result, V_2 and the speaker-effect decreased:

Therefore,

(1) the effect of C became the largest in the region between C and V_2 among all effects:

(2) the relative value of effect V_1 to other effects increased at part V_1.

This reveals that intensive correlation is seen between V_2 and speaker-effect.

(The correlation (canonical correlation) between vectors x and y has close relations with the regression theory between x and y.)

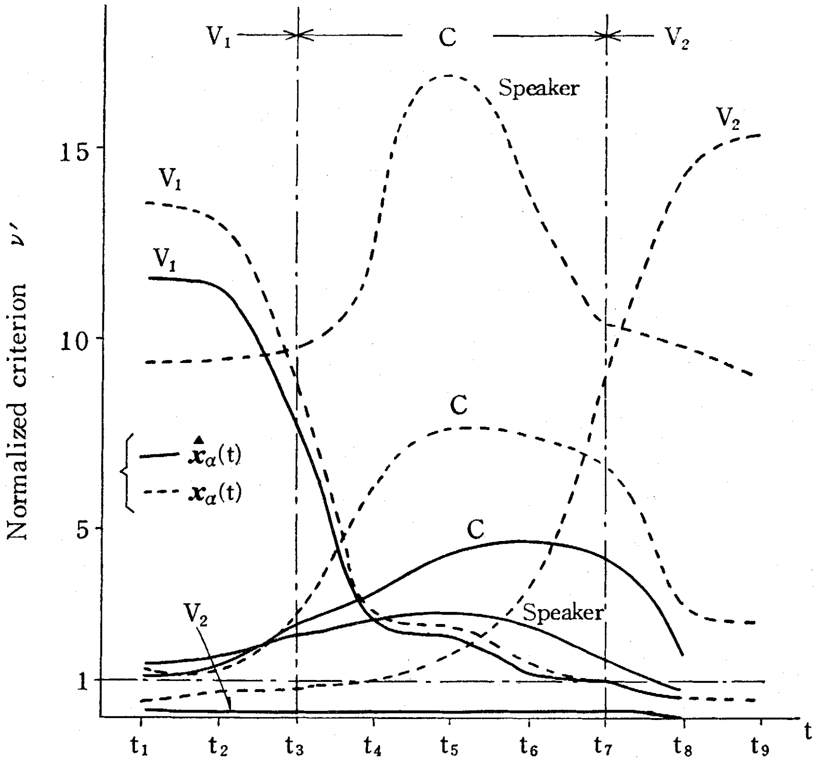

Fig. 10. Comparison between the analysis of variance of $\overset{\blacktriangle}{x}_\alpha$ and that of x_α. $\overset{\blacktriangle}{x}_\alpha$ is the difference between x_α and its regression estimate by $x_\alpha(t_9)$.

10 Conclusion

Considering the spectrum components as the components of multi-dimensional vector, we performed multivariate analysis of variance in sections at various time point of the $V_1 CV_2$ utterance with four factors—V_1, C, V_2 and speaker; and compared the amounts of the effects of each factor within themselves. Then we inspected the relation of the variance ellipsoids of each factor along their principal axes; signified that notion of direction as well as amount is necessary for explaning the effects of each factor; and compared these analyses with the principal-component analysis. Another thing we investigated by the method of regression estimate was the relation between final vowels and each section of words. Furthermore, we performed similar analysis on the basis of three-dimensional vectors which consist of the formant frequencies extracted from the same materials as above, and compared these with the case of spectra. The results concerning speech sounds are as follows:

(1) Speaker-effect is considerably large, while consonant-effect is not so large. However, the directions of three distributions of these two effects and vowel-effect meet at nearly right angles with each other:

(2) Intensive correlation is seen between vowel and speaker-factor:

(3) In the case of formant frequency, the informations on any factor other than vowel-factor are being decreased as compared with the case of spectrum distribution.

REFERENCES

1. D.J. Broad and R.H. Fertig, "Formant-Frequency Trajectories in Selected CVC-Syllable Nuclei", J. Acoust. Soc. Amer., 47, 6, p. 1572 (1970).
2. K. Tabata and T. Sakai, "Multivariate Analysis of VCV Syllables", Record of Joint Convention of the Acous. Soc. of Japan (Nov. 1971).
3. T. Sakai and K. Tabata, "Multivariate Statistical Analysis of VCV Syllables", Jour. of the Institute of Electronics and Communication Engineers of Japan 56-D, p. 63 (Jan. 1973).
4. T.W. Anderson, "An Introduction to Multivariate Statistical Analysis", John Wiley & Sons, New York (1958).
5. D.F. Morrison, "Multivariate Statistical Methods", McGraw-Hill, New York (1967).
6. M. Siotani and T. Asano, "Multivariate Analysis Theory", Kyoritsu, Tokyo (1967).
7. W. Klein, R. Plomp and L.C.W. Pols, "Vowel Spectra, Vowel Spaces, and Vowel Identification", J. Acoust. Soc. Amer., 48, 4, p. 999 (1970).
8. M.G. Kendall, "A Course in Multivariate Analysis", Charles Griffin, London (1957).
9. K. Ito, "Theory of Multivariate Statistical Analysis", p. 162, Baihu-kan, Tokyo (1969).
10. M. Nakatsui and J. Suzuki, "Formant Frequency Extraction by Inverse Filtering and Moment Calculation and Its Evaluation using Synthetic Speech Sound", J. Acoust. Soc. Japan, 26, 5, p. 211 (Jun. 1970).
11. S. Kitazawa, "A Quantitative Approach to Co-Articulation by Formant Frequencies", Graduation thesis for the bachelar degree of Kyoto University (1971).
12. S. Siegel, "Nonparametric Statistics for the Behavioral Sciences", p. 47, McGraw-Hill, New York (1956).

(Aug. 31, 1973, received)

Appndix A

The Asymptotic Distribution of the Likelihood Ratio Test Criterion in Multivariate Analysis of Variance for Four-Factor Design with Single Observation[(4)(5)(6)]

Let us consider the asymptotic distribution of the likelihood ratio test criterion indicated by Eq. (7). Arrange $\boldsymbol{x}_{ijkl}$ (the number of which is $n=a\cdot b\cdot c\cdot d$) of Eq. (2) in a certain order, and assume α-th vector to be $\boldsymbol{x}_\alpha$ ($\alpha=1,\ldots,n$), and $X'(p\times n)=(\boldsymbol{x}_1',\ldots,\boldsymbol{x}_n')$. Then let $F'(a\times n)=[f_{i\alpha}]$, $G'(b\times n)=[g_{j\alpha}]$, $H'(c\times n)=[h_{k\alpha}]$, and $M'(d\times n)=[m_{l\alpha}]$, where,

$$f_{i\alpha}=\begin{cases}1, & \text{if } \boldsymbol{x}_\alpha \in \{\boldsymbol{x}_{ijkl} \mid j=1,\ldots,b;\ k=1,\ldots,c;\ l=1,\ldots,d\}\\ 0, & \text{otherwise,}\end{cases}$$

and $g_{j\alpha}$, $h_{k\alpha}$, $m_{l\alpha}$ are similarly defined.

Meanwhile, if we define that $\mathbf{1}_n'(1\times n)=(1,\ldots,1)$, $B_1'(p\times a)=(\alpha_1',\ldots,\alpha_a')$, $B_2'(p\times b)=(\beta_1',\ldots,\beta_b')$, $B_3'(p\times c)=(\gamma_1',\ldots,\gamma_c')$ and $B_4'(p\times d)=(\delta_1',\ldots,\delta_d')$, it is possible to alter Eq. (2) to

$$\epsilon(X)=\mathbf{1}_n\mu+FB_1+GB_2+HB_3+MB_4=(\mathbf{1}_n\ FGHM)\,(\mu' B_1' B_2' B_3' B_4')'\equiv ZB$$

where ϵ denotes "expected value"; $Z=(\mathbf{1}_n\ FGHM)$ the design matrix of the experiment; B unknown parameter matrix. From Eq. (4), the density function of X is

$$(2\pi)^{-\frac{pn}{2}}|\Lambda|^{-\frac{n}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}\ \Lambda^{-1}(X-ZB)'(X-ZB)\right\},$$

where tr indicates the trace of matrix. From this density, we will find $\hat{B}$ and $\hat{\Lambda}$—the maximum likelihood estimates of B and Λ, respectively. B satisfies

$$Z'Z\hat{B}=Z'X, \tag{A.1}$$

but since $Z'Z$ has not its inverse matrix, we obtain $\hat{B}$ by comparing the both sides of (A.1) under the condition of Eq. (3), that is, $\mathbf{1}_n'B_1=0,\ldots,\mathbf{1}_n'B_4=0$.

$$\hat{\mu}=\boldsymbol{x}_{....},\ \hat{B}_1'=(\hat{\boldsymbol{\alpha}}_1',\ldots,\hat{\boldsymbol{\alpha}}_a)=(\boldsymbol{x}_{1...}'-\boldsymbol{x}_{....}',\ldots,\boldsymbol{x}_{a...}',-\boldsymbol{x}_{....}'),\ldots,\hat{B}_4'. \tag{A.2}$$

Therefore, the maximum likelihood estimate of Λ is

$$\begin{aligned}n\hat{\Lambda}&=(X-Z\hat{B})'(X-Z\hat{B})=(X-\mathbf{1}_n\hat{\mu}-F\hat{B}_1-G\hat{B}_2-H\hat{B}_3-M\hat{B}_4)'\\&\quad\cdot(X-\mathbf{1}_n\hat{\mu}-F\hat{B}_1-G\hat{B}_2-H\hat{B}_3-M\hat{B}_4)\\&=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i...}-\boldsymbol{x}_{.j..}-\boldsymbol{x}_{..k.}-\boldsymbol{x}_{...l}+3\boldsymbol{x}_{....})'\\&\qquad\cdot(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i...}-\boldsymbol{x}_{.j..}-\boldsymbol{x}_{..k.}-\boldsymbol{x}_{...l}+3\boldsymbol{x}_{....})\end{aligned} \tag{A.3}$$

which coincides with the residual R itself of Eq. (5).

Subsequently, consider the hypothesis that there is no effect of factor A

$$H_A:\ \boldsymbol{\alpha}_1=\ldots=\boldsymbol{\alpha}_a=0,\ \text{that is, } B_1=0 \tag{A.4}$$

Let $\hat{\hat{\mu}}$, $\hat{\hat{B}}_2$, $\hat{\hat{B}}_3$, $\hat{\hat{B}}_4$, and $\hat{\hat{\Lambda}}$ be the maximum likelihood estimates of μ, B_2, B_3, B_4, and Λ under the hypothesis H_A, then,

$$\hat{\hat{\mu}}=\boldsymbol{x}_{....}',\ \hat{\hat{B}}_2'=(\boldsymbol{x}_{.1..}'-\boldsymbol{x}_{....}',\ldots,\boldsymbol{x}_{.b..}'-\boldsymbol{x}_{....}'),\ \ldots,\ \hat{\hat{B}}_4',$$

$$n\hat{\hat{\Lambda}}=(X-1_n\hat{\mu}-G\hat{\hat{B}}_2-H\hat{\hat{B}}_3-M\hat{\hat{B}}_4)'(X-1_n\hat{\mu}-G\hat{\hat{B}}_2-H\hat{\hat{B}}_3-M\hat{\hat{B}}_4)$$
$$=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(x_{ijkl}-x_{\cdot j\cdot\cdot}-x_{\cdot\cdot k\cdot}-x_{\cdot\cdot\cdot l}+2x_{\cdot\cdot\cdot\cdot})'(x_{ijkl}-x_{\cdot j\cdot\cdot}$$
$$-x_{\cdot\cdot k\cdot}-x_{\cdot\cdot\cdot l}+2x_{\cdot\cdot\cdot\cdot})=bcd\sum_{i=1}^{a}(x_{i\cdot\cdot\cdot}-x_{\cdot\cdot\cdot\cdot})'(x_{i\cdot\cdot\cdot}-x_{\cdot\cdot\cdot\cdot})$$
$$+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(x_{ijkl}-x_{i\cdot\cdot\cdot}-x_{\cdot j\cdot\cdot}-x_{\cdot\cdot k\cdot}-x_{\cdot\cdot\cdot l}+3x_{\cdot\cdot\cdot\cdot})'$$
$$\cdot(x_{ijkl}-x_{i\cdot\cdot\cdot}-x_{\cdot j\cdot\cdot}-x_{\cdot\cdot k\cdot}-x_{\cdot\cdot\cdot l}+3x_{\cdot\cdot\cdot\cdot})=Q_1+R \tag{A.5}$$

Thus, the likelihood ratio criterion for the hypothesis H_A is

$$w=\frac{|n\hat{\Lambda}|}{|n\hat{\hat{\Lambda}}|}=\frac{|R|}{|Q_1+R|}. \tag{A.6}$$

This distribution of w under the hypothesis (A.4) is obtained from Cochran's theorem.

Let $Y=X-ZB=X-1_n\mu-FB_1-GB_2-HB_3-MB_4$, and break down $Y'Y$ as follows:

$$Y'Y=Y'\left(\frac{1_n1_n'}{n}\right)Y+Y'\left(\frac{FF'}{bcd}-\frac{1_n1_n'}{n}\right)Y+Y'\left(\frac{GG'}{acd}-\frac{1_n1_n'}{n}\right)Y+Y'$$
$$\cdot\left(\frac{HH'}{abd}-\frac{1_n1_n'}{n}\right)Y+Y'\left(\frac{MM'}{abc}-\frac{1_n1_n'}{n}\right)Y+Y'\left(I_n-\frac{FF'}{bcd}\right.$$
$$\left.-\frac{GG'}{acd}-\frac{HH'}{abd}-\frac{MM'}{abc}+3\frac{1_n1_n'}{n}\right)Y$$
$$\equiv Y'WoY+Y'W_1Y+Y'W_2Y+Y'W_3Y+Y'W_4Y+Y'W_5Y$$
$$\equiv V_0+V_1+V_2+V_3+V_4+V_5.$$

(I_n: n-dimensional unit matrix.)

Where $n\times n$ matrices—$W_0,\dots, W_5$—are all idempotent matrices, for example, as

$$W_1^2=\left(\frac{FF'}{bcd}-\frac{1_n1_n'}{n}\right)^2=\frac{FF'}{bcd}-\frac{1_n1_n'}{n}=W_1.$$

These are easily verified if we pay attention to following relations.

$$F'F=bcd\ I_a,\quad 1_n'F=bcd\ 1_a',\quad F1_a=1_n,\quad F'G=cd\ 1_a1_b'.$$

Therefore, taking the trace of matrices in order to obtain the ranks of each matrix,

$$tr(W_0)=1,\ tr(W_1)=\frac{1}{bcd}\sum_{\alpha=1}^{n}\sum_{i=1}^{a}f_{i\alpha}^2-\frac{n}{n}=\frac{n}{bcd}-1=a-1,$$
$$tr(W_2)=b-1,\quad tr(W_3)=c-1,\quad tr(W_4)=d-1,$$
$$tr(W_5)=n-a-b-c-d+3 \tag{A.7}$$

are obtained. And since $n=(1)+(a-1)+(b-1)+(c-1)+(d-1)+(n-a-b-c-d+3)$, $V_0,\dots, V_5$ are independent of each other and can be expressed by Cochran's theorem as follows:

$$V_i=\sum_{j=1}^{\tau_i}u_j'u_j,\qquad u_j(1\times p)\sim N(0,\Lambda),\ (i=0,1,\dots,5).$$

Where, τ_i is the rank of W_i shown in Eq. (A.7), and u_j are assumed to be distributed independently of each other. (When $\tau_i>p$, V_i is distributed according to Wishart distribution $W(\Lambda, p, \tau_i)$.) Now, keeping attention to $1_a'B_1=1_b'B_2=1_c'B_3=1_d'B_4=0$, substitute $X-ZB$ for Y in V_1, and express V_1 as the function of x_{ijkl}. Then

$$V_1=bcd\sum_{i=1}^{a}(\boldsymbol{x}_{i}\ldots-\boldsymbol{x}\ldots.-\boldsymbol{\alpha}_i)'(\boldsymbol{x}_{i}\ldots-\boldsymbol{x}\ldots.-\boldsymbol{\alpha}_i).$$

Particularly when $H_A: B_1=0$ is true, $V_1=Q_1$.

On the other hand,

$$V_5=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i}\ldots-\boldsymbol{x}_{\cdot j}\ldots-\boldsymbol{x}_{\cdot\cdot k\cdot}-\boldsymbol{x}_{\ldots l}+3\boldsymbol{x}\ldots.)'$$
$$\cdot(\boldsymbol{x}_{ijkl}-\boldsymbol{x}_{i}\ldots-\boldsymbol{x}_{\cdot j}\ldots-\boldsymbol{x}_{\cdot\cdot k\cdot}-\boldsymbol{x}_{\ldots l}+3\boldsymbol{x}\ldots.)=R \qquad (A.8)$$

$V_5=R$ holds no matter what $B_1\sim B_4$ may be, and $R\sim W(\Lambda, p, \tau_5)$ when $\tau_5=n-a-b-c-d+3\geqq p$.

Therefore, the moment of w of Eq. (A.6) is accurately obtained so that we can learn its asymptotic distribution by Box.

Namely, $\nu=-\{n-l_2-(p+l_1+1)/2\}\log w$ is distributed as a Chi-squared variate with pl_1 degrees of freedom as the sample size n tends to infinity. Where $l_1=a-1$, $l_2=b+c+d-2$, $n-a-b-c-d+3\geqq p$.

Appendix B

The Intersection Produced by the Ellipsoid and the Straight Line or the Plane

The point of intersection, produced by the ellipsoid $\boldsymbol{x}A^{-1}\boldsymbol{x}'=1$ and the straight line kC through the origin (k represents arbitrary real number), is expressed by $\pm C/\sqrt{CA^{-1}C'}$. (It is obtained by substituting kC for $\boldsymbol{x}$ in $\boldsymbol{x}A^{-1}\boldsymbol{x}'=1$.)

As the distance between the origin and the point of intersection is

$$\sqrt{\frac{CC'}{CA^{-1}C'}} \qquad (B.1)$$

this expression coincides with $\sqrt{\lambda}$ when C is the eigenvector of $AC'=\lambda C'$. ($C'=\lambda A^{-1}C'$; $CC'=\lambda CA^{-1}C$; $\lambda=CC'/CA^{-1}C'$.)

The intersection, drawn by the ellipsoid and the plane determined by C_1 and C_2, will be obtained by connecting the points of intersections which are produced by the ellipsoid and arbitrary straight lines on the plane which pass through the origin.

Since an arbitrary straight line is expressed by $k_1C_1+k_2C_2$ upon choosing k_1 and k_2 arbitrarily, the distance between the point of intersection (made by the straight line) and the origin is obtained by substituting $k_1C_1+k_2C_2$ for C in Eq. (B.1).

APPENDIX C

Variance of Factor A

Let $\epsilon_A(x_{ijkl})$ be the expected value of x_{ijkl} $(=\mu+\alpha_i+\beta_j+\gamma_k+\delta_l+\varepsilon_{ijkl})$ under the hypothesis that $H_A: \alpha_1=\ldots=\alpha_a=0$ is true. Then, from Eq. (A.5) in Appendix A, the maximum likelihood estimate of $\epsilon_A(x_{ijkl})$ is

$$\hat{\hat{\epsilon}}_A(x_{ijkl})=\hat{\hat{\mu}}+\hat{\hat{\beta}}_j+\hat{\hat{\gamma}}_k+\hat{\hat{\delta}}_l=x_{\cdots\cdot}+(x_{\cdot j\cdot\cdot}-x_{\cdots\cdot})+(x_{\cdot\cdot k\cdot}-x_{\cdots\cdot})+(x_{\cdots l}-x_{\cdots\cdot}) \tag{C.1}$$

Since $\mu-\hat{\hat{\mu}}\cong 0$, $\beta_j-\hat{\hat{\beta}}_j\cong 0$, $\gamma_k-\hat{\hat{\gamma}}_k\cong 0$ and $\delta_l-\hat{\hat{\delta}}_l\cong 0$,
the difference between sample vector x_{ijkl} and $\hat{\hat{\epsilon}}_A(x_{ijkl})$ is approximately equal to $\alpha_i+\varepsilon_{ijkl}$ as follows:

$$x_{ijkl}-\hat{\hat{\epsilon}}_A(x_{ijkl})=(\mu-\hat{\hat{\mu}})+\alpha_i+(\beta_j-\hat{\hat{\beta}}_j)+(\gamma_k-\hat{\hat{\gamma}}_k)+(\delta_l-\hat{\hat{\delta}}_l)+\varepsilon_{ijkl}\cong\alpha_i+\varepsilon_{ijkl} \tag{C.2}$$

Therefore, the above difference may be considered to be the deviation only due to factor A $(A_1,\ldots,A_a)$.
(Of course, if hypothesis H_A is actually true (it means $\alpha_i=0$), $x_{ijkl}-\hat{\hat{\epsilon}}_A(x_{ijkl})\cong\varepsilon_{ijkl}$ and this difference is the residual itself.) Hence the matrix of sums of squares and cross products on the basis of the deviation (C.2)

$$\begin{aligned}
&\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(x_{ijkl}-\hat{\hat{\epsilon}}_A(x_{ijkl}))'(x_{ijkl}-\hat{\hat{\epsilon}}_A(x_{ijkl}))\\
&=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(x_{ijkl}-x_{\cdots\cdot}-(x_{\cdot j\cdot\cdot}-x_{\cdots\cdot})-(x_{\cdot\cdot k\cdot}-x_{\cdots\cdot})\\
&-(x_{\cdots l}-x_{\cdots\cdot}))'(x_{ijkl}-x_{\cdots\cdot}-(x_{\cdot j\cdot\cdot}-x_{\cdots\cdot})-(x_{\cdot\cdot k\cdot}-x_{\cdots\cdot})\\
&-(x_{\cdots l}-x_{\cdots\cdot}))=\mathrm{bcd}\sum_{i=1}^{a}(x_{i\cdots}-x_{\cdots\cdot})'(x_{i\cdots}-x_{\cdots\cdot})\\
&+\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{d}(x_{ijkl}-x_{i\cdots}-x_{\cdot j\cdot\cdot}-x_{\cdot\cdot k\cdot}-x_{\cdots l}+3x_{\cdots\cdot})\\
&\cdot(x_{ijkl}-x_{i\cdots}-x_{\cdot\cdot j\cdot}-x_{\cdot\cdot k\cdot}-x_{\cdots l}+3x_{\cdots\cdot})=Q_1+R
\end{aligned} \tag{C.3}$$

can be considered to represent the variance of factor A.

Moreover, let $\hat{\hat{\Lambda}}$ be the maximum likelihood estimate of the covariance matrix Λ under the hypothesis $H_A: \alpha_1=\ldots=\alpha_a$. Then, from Eq. (A.5) in Appendix A

$$n\hat{\hat{\Lambda}}=Q_1+R,$$

whose expression is coincident with Eq. (C.3).

APPENDIX D

The Vector Space Normalized by the Residual R

Suppose the following nonsingular transformation (not always orthogonal transformation).

$$\boldsymbol{x} \rightarrow \tilde{\boldsymbol{x}} = \boldsymbol{x}(\mathrm{R}/\mathrm{n})^{-\frac{1}{2}} \tag{D.1}$$

Where R is the residual given by Eq. (5), n is the sample size (n=abcd), and $\boldsymbol{x}$ and $\tilde{\boldsymbol{x}}$ are vectors in the original and new spaces, respectively.

Then, $\frac{1}{\mathrm{n}}\mathrm{Q}_1 = \frac{1}{\mathrm{n}}\sum_{i=1}^{a}(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)'(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)$ (Q_1 given by Eq. (5)) is transformed to

$$\begin{aligned}\frac{1}{\mathrm{n}}\tilde{\mathrm{Q}}_1 &= \frac{1}{\mathrm{n}}\sum_{i=1}^{a}(\tilde{\boldsymbol{x}}_{i}\ldots - \tilde{\boldsymbol{x}}\ldots.)'(\tilde{\boldsymbol{x}}_{i}\ldots - \tilde{\boldsymbol{x}}\ldots.)\\ &= \frac{1}{\mathrm{n}}\sum_{i=1}^{a}(\mathrm{R}/\mathrm{n})^{-\frac{1}{2}}(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)'(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)(\mathrm{R}/\mathrm{n})^{-\frac{1}{2}}\\ &= \mathrm{R}^{-\frac{1}{2}}\left\{\sum_{i=1}^{a}(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)'(\boldsymbol{x}_{i}\ldots - \boldsymbol{x}\ldots.)\right\}\mathrm{R}^{-\frac{1}{2}}\\ &= \mathrm{R}^{-\frac{1}{2}}\mathrm{Q}_1\mathrm{R}^{-\frac{1}{2}}\end{aligned} \tag{D.2}$$

and

$$\frac{1}{\mathrm{n}}\mathrm{R}(\mathrm{p}\times\mathrm{p}) \rightarrow \frac{1}{\mathrm{n}}\tilde{\mathrm{R}} = \mathrm{R}^{-\frac{1}{2}}\mathrm{R}\mathrm{R}^{-\frac{1}{2}} = \mathrm{I}_\mathrm{p}. \tag{D.3}$$

Similarly,

$$\frac{1}{\mathrm{n}}(\mathrm{Q}_1+\mathrm{R}) \rightarrow \frac{1}{\mathrm{n}}(\tilde{\mathrm{Q}}_1+\tilde{\mathrm{R}}) = \mathrm{R}^{-\frac{1}{2}}(\mathrm{Q}_1+\mathrm{R})\mathrm{R}^{-\frac{1}{2}}, \tag{D.4}$$

and

$$\begin{aligned}&\tilde{\boldsymbol{x}}\left[\frac{1}{\mathrm{n}}(\tilde{\mathrm{Q}}_1+\tilde{\mathrm{R}})\right]^{-1}\tilde{\boldsymbol{x}}' = \tilde{\boldsymbol{x}}\,[\mathrm{R}^{-\frac{1}{2}}(\mathrm{Q}_1+\mathrm{R})\mathrm{R}^{-\frac{1}{2}}]^{-1}\tilde{\boldsymbol{x}}' = \tilde{\boldsymbol{x}}(\mathrm{R}/\mathrm{n})^{\frac{1}{2}}\\ &\cdot\left[\frac{1}{\mathrm{n}}(\mathrm{Q}_1+\mathrm{R})\right]^{-1}(\mathrm{R}/\mathrm{n})^{\frac{1}{2}}\,\tilde{\boldsymbol{x}}' = \boldsymbol{x}\left(\frac{1}{\mathrm{n}}(\mathrm{Q}_1+\mathrm{R})\right)^{-1}\boldsymbol{x}'.\end{aligned} \tag{D.5}$$

APPENDIX E

The Kolmogorov-Smirnov One-Sample Test[12]

The Kolmogorov-Smirnov one-sample test is a test of goodness of fit. That is, it is concerned with the degree of agreement between the distribution of a set of sample values (observed scores) and some specified theoretical distribution. It determines whether the scores in the sample can reasonably be thought to have come from a population having the theoretical distribution.

Briefly, the test involves specifying the cumulative frequency distribution which would occur under the theoretical distribution and comparing that with the observed cumulative frequency distribution. The theoretical distribution represents what would be expected under H_0.

Let $F_0(X)$=a completely specified cumulative frequency distribution function, the theoretical cumulative distribution under H_0. That is, for any value of X, the value of $F_0(X)$ is the proportion of cases expected to have scores equal to or less than X.

And let $S_N(X)$=the observed cumulative frequency distribution of a random sample of N observations. Where X is any possible score, $S_N(X)=k/N$, where k=the number of observations equal to or less than X.

Now under the null hypothesis that the sample has been drawn from the specified theoretical distribution, it is expected that for every value of X, $S_N(X)$ should be fairly close to $F_0(X)$. That is, under H_0 we would expect the differences between $S_N(X)$ and $F_0(X)$ to be small and within the limits of random errors. The Kolmogorov-Smirnov test focuses on the largest of the deviations. The largest value of $F_0(X)-S_N(X)$ is called the maximum deviation, D;

$$D=\text{maximum } |F_0(X)-S_N(X)| \qquad \text{(E.1)}$$

The sampling distribution of D under H_0 is known. The Table E.1 gives D_α, the critical value of D when N is over 35. If N is over 35, one determines D_α, the critical values of D by the divisions indicated in Table E.1. For example, suppose a researcher uses N=43 cases and sets α=.05. Table (E.1) shows that any D equal to or greater than $D_\alpha=\frac{1.36}{\sqrt{N}}$ will be significant. That is, any D, as defined by formula (E.1), which is equal to or greater than $D_\alpha=\frac{1.36}{\sqrt{43}}=.207$ will be significant at the .05 level (two-tailed test).

Table E.1. Table of D_α (critical value of D) in the Kolmogorov-Smirnov one-sample test.

Sample size (N)	α, Level of signigicance for $D=\max|F_0(X)-S_N(X)|$				
	0.20	0.15	0.10	0.05	0.01
Over 35	$\frac{1.07}{\sqrt{N}}$	$\frac{1.14}{\sqrt{N}}$	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$